DFW Housing Market

Brandilyn Hall, Elyssa Irizarry, Megan Adams, Yvonne Martinez

Agenda

About our Data
- Topic & Source
- Research Questions & Hypothesis

Machine Learning
- Data Processing & Target Selection
- Modeling & Model Selection

Web App & Tableau
- Tour & Demonstration
- Discoveries

Conclusions
- Final Words & Future Work
- Questions

Overview

- We used DFW housing data to predict the closing price of potential sellers' houses which we built within our web app

- We also utilized Tableau to further visualize our housing data and provide information about the demographics of the DFW area

Our Data

DFW Area Housing Data
NTREIS (North Texas Real Estate Information Service)

Demographic Information
US Census (for Tableau purposes)

Inspiration

- Texas Quarterly Housing Report (visualizations)
- TAMU Interactive Housing Dashboard
- Kaggle
  - Austin Housing EDA
  - House Price Prediction
- Tableau Public

Theme and Design

- Went off of NTREIS logo
- Many real estate ads and marketing material have either red or blue incorporated

Research Questions & Hypothesis

- Which area of DFW has the fastest sale rate?
- Which area of DFW sells houses for the most money?
- Which features attribute to making a house sell faster and for more money?

Hypothesis: We propose that location, square footage, number of bed/bathrooms, and additions (e.g, garage, pools, sheds) affects the price.

Machine Learning

- Data Processing
- Feature Selection
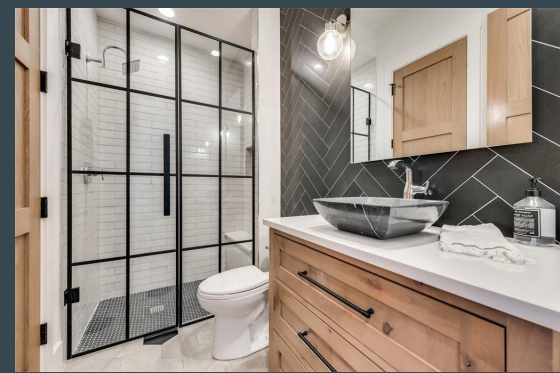- Modeling
- Model Selection:

Website Demo

House Showcase

511 Coopers Bluff
(9 bathrooms)

Heath, TX

Coopers Bluff Bathrooms

1289 Viento Oaks
(most covered parking - 9)

Fort Worth, TX

1806 High Country Drive (most expensive $5.6 mil)
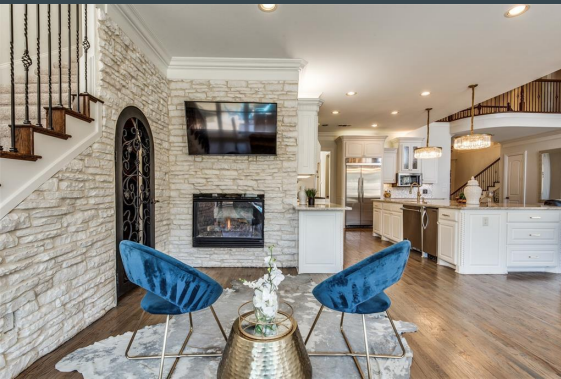
Westlake, TX

3508 Greystone Court (6 fireplaces)

McKinney, TX

Greystone Court Fireplaces

3932 Potomac
Most Expensive ($14 mil)

Highland Park TX

3932 Potomac

Conclusions

- Features such as rooms, garages, etc affect the price
  - The more features, the higher the price
- Location affects price (affluent neighborhoods)
- Houses with no HOA's are actually **more** expensive

Limitations

- Limited Dataset
  - Limited to DFW
  - Last 5,000 rows in the last 6 months
- Days on Market not able to predict

Future Work

- Conduct similar analysis on various areas of the US
- Obtain data to successfully predict days on market
- Setup automatic feed of new data to obtain current and accurate predictions

Questions?